A MULTIRATE APPROACH TO HIGH-RESOLUTION IMAGE RECONSTRUCTION

James W. Scrofani
Department of Electrical and Computer Engineering
Naval Postgraduate School
Monterey, California 93943-5000
email: jwscrofa@nps.edu

Charles W. Therrien
Department of Electrical and Computer Engineering
Naval Postgraduate School
Monterey, California 93943-5000
email: therrien@nps.edu

ABSTRACT

The paper addresses the problem of reconstructing a high resolution image from a set of observation images sampled at a lower rate and subject to additive noise and distortion. The method introduced here is based on our work in multirate optimal filtering extended to two dimensions. The linear filters used for the reconstruction are periodically spatially-varying (in 2-D) and chosen so that their region of support are closest to the point being estimated. Results are presented for images with additive white noise and compared to methods using interpolation.

KEY WORDS

Super Resolution, Image Reconstruction, Weiner-Hopf, Optimal Filtering, Multirate, Stochastic.

1 Introduction

Super resolution (SR) imaging has recently become an area of great interest in the image processing research community [1, 2, 3]. The ability to form a high resolution (HR) image from a collection of subsampled images has a broad range of applications and has largely been motivated by physical and production limitations on existing image acquisition systems, and the marginal costs associated with increase spatial resolution. Figure (1) depicts the SR concept, where a collection of LR images of a scene are superimposed on a HR grid, available for subsequent HR image reconstruction.

In this paper we propose a stochastic multirate approach to this problem, adapting and extending the work in [4, 5, 6, 7, 8] to two-dimensional signals. The earlier work has focused on information fusion, i.e., on the combination of observations from multiple sensors to perform tracking, surveillance, classification or some other task, and on the reconstruction of one-dimensional signals from multiple observations at a lower rate. This work extends these concepts to SR image reconstruction.

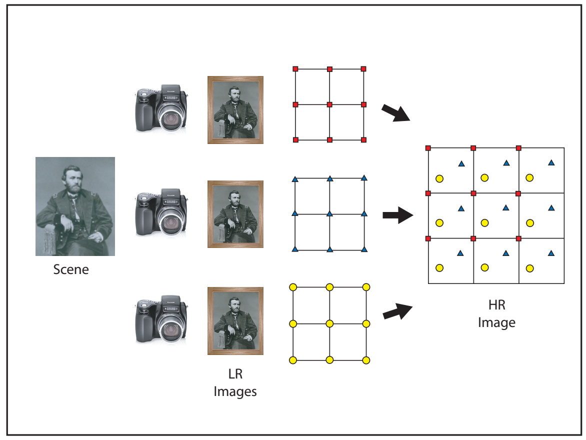

Figure 1. The super resolution imaging concept.

2 Proposed Method

2.1 Observation Model

The relationship between the set of low resolution observations and the underlying high resolution image is modeled in the block diagram of Figure (2). This sensor model shows that each LR observation is acquired from the HR image subject to distortion by linear filtering (typically blur), subpixel translation, downsampling, and channel noise. The parameters L_1 and L_2 represent the horizontal and vertical downsampling factors, respectively. While, the parameters i and j represent the horizontal and vertical subpixel translation, respectively. The parameter $\boldsymbol{N}_{ij}$ represents additive white Gaussian noise. We can represent the observation model as

$$\boldsymbol{F}_{ij} = \boldsymbol{D}_{L_1}^{(i)} \boldsymbol{F} \boldsymbol{G}_{ij} \boldsymbol{D}_{L_2}^{(j)T} + \boldsymbol{N}_{ij} \qquad (1)$$

The matrix $\mathbf{D}_L^{(k)}$ is called a "decimation matrix with time delay" [7] and is used to extract the appropriate pixels to form each observation matrix. The matrix is defined in terms of a Kronecker product of the form

$$\mathbf{D}_L^{(k)} = \mathbf{I} \otimes \boldsymbol{\iota}_k \qquad 0 \leq k \leq L-1 \qquad (2)$$

where $\mathbf{I}$ is the $P_i \times P_i$ identity matrix and $\boldsymbol{\iota}_k$ is a $1 \times L$ index vector with a 1 in the $k+1^{th}$ position and 0's elsewhere.

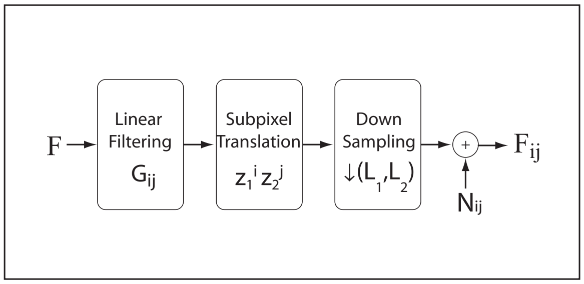

Figure 2. Observation model relating HR image with LR observations.

2.2 Estimate

Consider the set of LR observations $\{\boldsymbol{F}_{ij}\}$ and the related HR image $\boldsymbol{F}$. We desire to form an estimate for the HR image by some weighted sum of the LR observations. The estimate can be written as

$$\hat{F}[n_1, n_2] = \sum_{i=0}^{L_1-1} \sum_{j=0}^{L_2-1} \langle \boldsymbol{f}_{ij}[n_1, n_2], \boldsymbol{H}_{ij}^{(k_1,k_2)} \rangle \tag{3}$$

where the expression on the right represents the Frobenius inner product of the matrices, defined by

$$\langle \boldsymbol{A}, \boldsymbol{B} \rangle \equiv \mathrm{tr}(\boldsymbol{A}\boldsymbol{B}^T) \equiv \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij},$$

where $\boldsymbol{A}$ and $\boldsymbol{B} \in \mathbb{R}^{m \times n}$. Here the matrix $\boldsymbol{f}_{ij}[n_1, n_2]$ of size $P \times Q$ is the set of image points lying within the appropriate LR image mask and matrix $\boldsymbol{H}_{ij}^{(k_1,k_2)}$ is the corresponding set of filter coefficients. The masks are chosen to minimize the mean-square error

$$\mathcal{E}\{\|\vec{\boldsymbol{F}} - \hat{\vec{\boldsymbol{F}}}\|_F^2\} \tag{4}$$

where subscript 'F' represents the Frobenius norm. The parameters $k_i = n_i \bmod L_i$ for $0 \leq k_i \leq L_i$ and (i, j) represent subpixel translation with $0 \leq i \leq L_1 - 1$, $0 \leq j \leq L_2 - 1$, and $(i, j) \in \mathbb{Z}$. In the maximally decimated case, i and j span the entire set $\{0, 1, \ldots, L_1 - 1\}$ and $\{0, 1, \ldots, L_1 - 1\}$, respectively. Both $\{\boldsymbol{f}_{ij}[n_1, n_2]\}$ and $\boldsymbol{H}_{ij}^{(k_1,k_2)}$ are further described in Sections 2.3 and 2.4, respectively.

2.3 Index Mapping/LR Image Mask

Developing the LR image masks involves mapping indices in the HR sample index domain $\mathbb{D}_1$ to those in the LR sample index domain $\mathbb{D}_2$. For a given pixel intensity $F[n_1, n_2]$, the HR indices $[n_1, n_2]$ map to a set of sample indices $\{[m_1, m_2]_{ij}\}$ which correspond to pixel intensities $\{F_{ij}[m_1, m_2]\}$. The indices corresponding to each observation matrix $\boldsymbol{F}_{ij}$ are determined such that

$$|[n_1, n_2] - [m_1, m_2]_{ij}| \tag{5}$$

is minimized for each (i, j). This mapping can be shown to be

$$[m_1, m_2]_{ij} = [T(n_1), T(n_2)]. \tag{6}$$

where $T[n]$ is defined by

$$T : n \rightarrow \frac{n + 1 - (k - i + 1) \bmod L - i}{L} \tag{7}$$

where $n \in \mathbb{D}_1$, $k = n \bmod L$, and i is the translation.

The LR indices $[m_1, m_2]_{ij}$ for each observation represent the centroid of each of the LR image masks. Given a desired mask of size $P \times Q$, each LR image mask is comprised of the $P \times Q$ pixels closest to $\boldsymbol{F}_{ij}[m_1, m_2]$.

2.4 Filter Mask

If the desired HR image $\boldsymbol{F}$ and its observations $\boldsymbol{F}_{ij}$ are jointly homogeneous, then the linear filters $\boldsymbol{H}_{ij}^{(k_1,k_2)}$required for optimal estimation are periodically spatially-varying, an extension of [5]. This periodicity can be described in terms of the "phase" (k_1, k_2), where $k_i = n_i \bmod L_i$. If we define the set of *least positive residues* as $\mathbb{Z}_n = \{0, 1, \ldots, n - 1\}$, then $k_1 \in \mathbb{Z}_{L_1}$ and $k_2 \in \mathbb{Z}_{L_2}$, and all possible configurations of phase can be represented as $\mathbb{Z}_{L_1} \times \mathbb{Z}_{L_2}$.

Figure (3) depicts the phase variation for $L_1 = L_2 = 2$. In this case $\mathbb{Z}_2 = \{0, 1\}$ and $\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0,0), (0,1), (1,0), (1,1)\}$. The spatial periodicity of the phase can be observed in this example by noting the regular recurrence of phase terms.

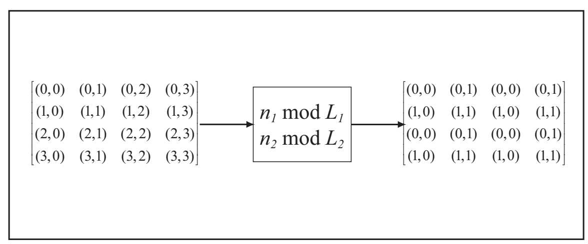

Figure 3. Index representation to modulo representation,with $L_1 = L_2 = 2$ (note the spatial phase periodicity).

2.5 Least Squares Formulation

In order to determine the filter coefficients required to estimate the HR image, a least squares (LS) approach is employed. We identify the set of all HR pixels that correspond to a given phase and denote this set of HR pixels as the matrix $\boldsymbol{F}^{(k_1,k_2)}$. This concept is depicted in Figure (4) where each shape corresponds to a unique phase (circle $(0,0)$, square $(0,1)$, triangle $(1,0)$, and star $(1,1)$). From

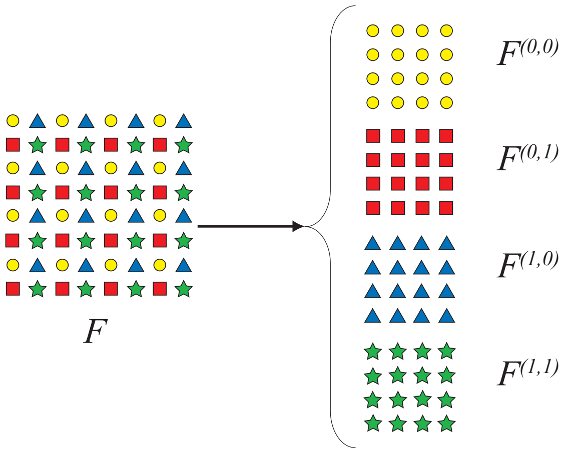

Figure 4. Relationship between HR pixels and spatially-varying filter masks in formulating the LS problem.

Equation (3) we can see that

$$\begin{bmatrix} F[l_1, l_2] \\ F[m_1, m_2] \\ F[n_1, n_2] \\ \vdots \end{bmatrix} = \begin{bmatrix} \sum\sum\langle \boldsymbol{f}_{ij}[l_1, l_2], \boldsymbol{H}_{ij}^{(k_1,k_2)} \rangle \\ \sum\sum\langle \boldsymbol{f}_{ij}[m_1, m_2], \boldsymbol{H}_{ij}^{(k_1,k_2)} \rangle \\ \sum\sum\langle \boldsymbol{f}_{ij}[n_1, n_2], \boldsymbol{H}_{ij}^{(k_1,k_2)} \rangle \\ \vdots \end{bmatrix} \quad (8)$$

where the left hand expression is $\vec{\boldsymbol{F}}^{(k_1,k_2)}$. From this we can write

$$\vec{\boldsymbol{F}}^{(k_1,k_2)} \stackrel{\text{ls}}{=} \boldsymbol{\Phi} \vec{\boldsymbol{H}}_{ij}^{(k_1,k_2)} \quad (9)$$

where $\boldsymbol{\Phi}$ is the data matrix. This system of equations is solved in a least squares sense for the required set of filter masks at a given phase $\{\boldsymbol{H}_{ij}^{(k_1,k_2)}\}$.

2.6 Processing Method

The process used in the SR image reconstruction of a set of LR images is described as follows. First, a HR training image is obtained that is representative of the class of images that will be processed. From this image, a maximally decimated set of LR observations are derived and then through the least squares methodology of Section 2.5, filter coefficients are computed. With the class-specific filter coefficients, we are able to reconstruct images “of the same class”, by employing the estimate of Equation (3). In other words, we use training data to develop the filter masks from a class representative HR image, then filter any sets of LR images that are members of this class to reconstruct HR images.

3 Results

In order to evaluate the performance of this method of image reconstruction, we process the skyline image depicted in Figure (5), subject to varying degrees of additive white Gaussian noise. The image used for the training process is the 204×204 pixel subimage depicted in this figure. From this image, a set of filter masks is derived that is used in SR image reconstruction.

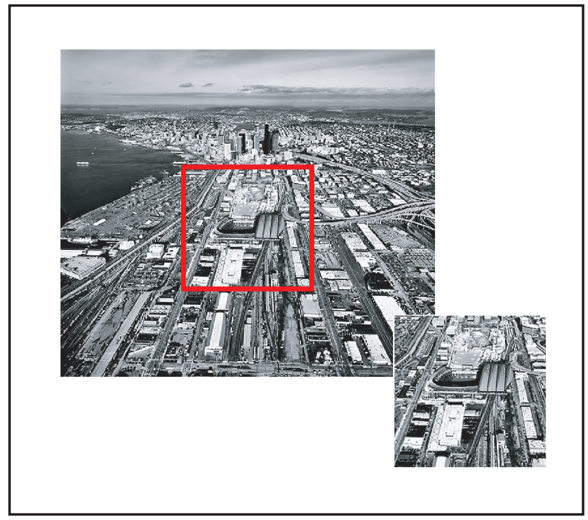

Figure 5. Subimage used to train filter.

The target or object of the reconstruction is depicted in Figure (6). From this 204×204 pixel subimage, LR observations are derived and are filtered using the class-specific filter masks. The same level of additive white Gaussian noise is used for the training and target images.

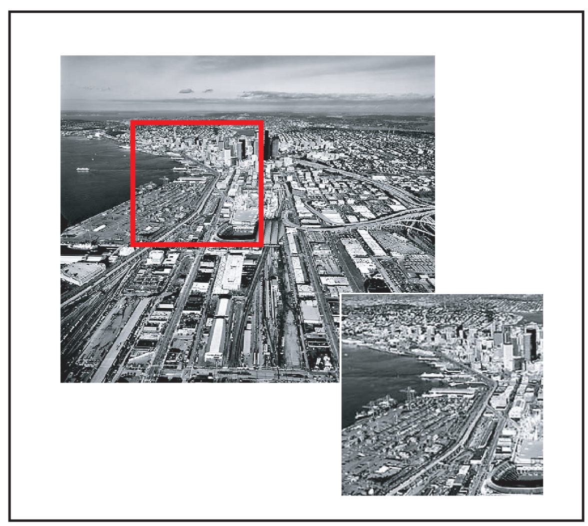

Figure 6. Subimage to be estimated.

Figure (7) depicts three members of the set of LR observations with various subpixel translations. The first image represents subpixel translation by one pixel in the horizontal direction and no translation in the vertical direction. The second represents translation by one pixel in both the vertical and horizontal directions. The third image represents translation by two pixels in both directions.

Figure 7. Downsampled observation images with subpixel translations $(1, 0)$, $(1, 1)$, and $(2, 2)$, respectively. $L = 3$, $P = Q = 3$, No AWGN.

In the remaining figures, the left panel depicts the SR image reconstruction using the proposed algorithm and the right panel depicts nearest-neighbor interpolation of one of the LR observations. In every case, the proposed method is superior to the interpolated result. During these experiments, other interpolation methods were considered, including bilinear and bicubic methods, but again the proposed method was superior.

Figure (8) compares reconstructed and interpolated images for the case of no additive noise, downsampling by 3 in both the vertical and horizontal directions, and with filter mask size of 3×3. In this case, the reconstruction yields a result that is visibly indiscernible from the target image.

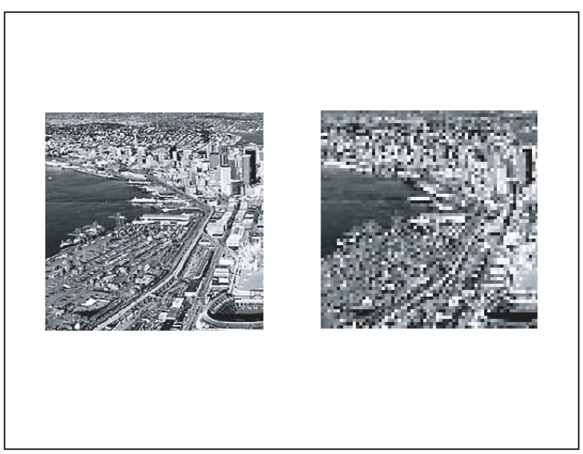

Figure 8. Comparison between a reconstructed image and interpolated image with $L = 3$, $P = Q = 3$, No AWGN.

Figure (9) compares reconstructed and interpolated images for the case of an SNR $= 5$ dB, downsampling by 3 in both the vertical and horizontal directions, and with filter mask size of 3×3. In this case, we see the effects of additive noise on the reconstruction. Despite some blurring of edges, details are still discernible.

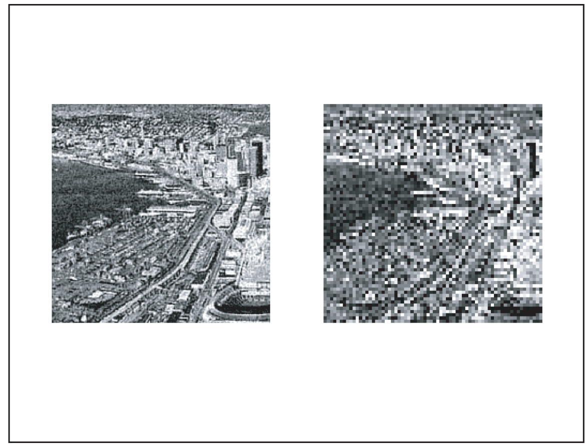

Figure 9. Comparison between a reconstructed image and interpolated image with $L = 3$, $P = Q = 3$, $SNR = 5$ dB.

Figure (10) compares reconstructed and interpolated images for the case of an SNR $= -1.5$ dB, downsampling by 3 in both the vertical and horizontal directions, and with filter mask size of 3×3. In this case, the effects of additive noise on the reconstruction are quite deleterious. Further blurring of edges is evident, and details have become hard to see. Major features in the image are still discernible however. Thus there is a significant advantage in using the proposed method over interpolation, where not even major features are discernible.

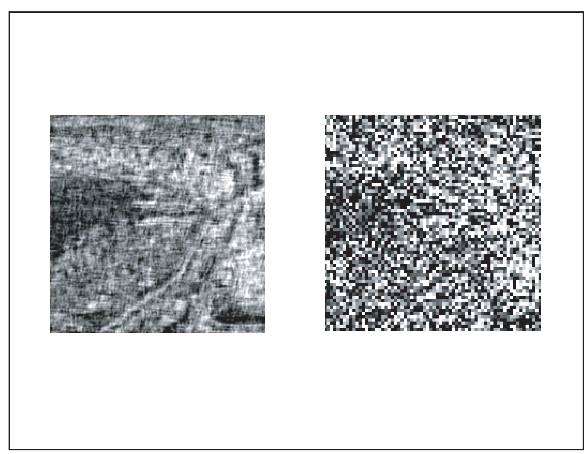

Figure 10. Comparison between a reconstructed image and interpolated image with $L = 3$, $P = Q = 3$, $SNR = -1.5$ dB.

4 Conclusion

In this paper a multirate optimal linear filtering approach is introduced for SR image reconstruction. Given some underlying HR image that cannot be observed directly, the set of related subsampled and translated LR images are combined, in an optimal fashion, to yield an estimate of the HR image. This method involves development of optimal periodically space-varying filters that are applied to these LR observations. In order to evaluate the performance of this method, several simulations were carried out, with varying degrees of additive white Gaussian noise. To evaluate the performance the results of this reconstruction method were compared to nearest-neighbor interpolation of LR observations. In every case, the proposed method performed better.

References

[1] S. Chaudhuri and R. D. Taur, "High-resolution slow-motion sequencing," *IEEE Signal Processing Magazine*, vol. 22, pp. 16 – 24, Mar. 2005.

[2] S. C. Park, M. K Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, pp. 21 – 36, May 2003.

[3] S. Chaudhuri, Ed., *Super-resolution Imaging*, Kluwer Academic Publishers, 2001.

[4] R. Cristi, D. A. Koupatsiaris, and C. W. Therrien, "Multirate filtering and estimation: The multirate Wiener filter," in *Proc.* 34^{th} *Asilomar Conf. on Signals, Systems, and Computers*, Oct. 2000, vol. 1, pp. 450–454.

[5] C. W. Therrien, "Issues in multirate statistical signal processing," in *Proc.* 35^{th} *Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2001, vol. 1, pp. 573–576.

[6] A. H. Hawes and C. W. Therrien, "Least squares optimal filtering with multirate observations," in *Proc.* 36^{th} *Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2002, vol. 2, pp. 1782–1786.

[7] R. J. Kuchler and C. W. Therrien, "Optimal filtering with multirate observations," in *Proc.* 37^{th} *Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2003, vol. 1, pp. 1208–1212.

[8] J. W. Scrofani and C. W. Therrien, "A stochastic multirate signal processing approach to high-resolution signal reconstruction," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 2005, Philadelphia.